Improving the Reproducibility of Diagnosing Micrometastases and Isolated Tumor Cells

Gábor Cserni, M.D., Ph.D.[1]
Simonetta Bianchi, M.D.[2]
Werner Boecker, M.D., Ph.D.[3]
Thomas Decker, M.D.[4]
Manuela Lacerda, M.D.[5]
Fritz Rank, M.D.[6]
Clive A. Wells, B.Chir.[7]
for The European Working Group for Breast Screening Pathology

[1] Department of Pathology, Bács-Kiskun County Teaching Hospital, Kecskemét, Hungary.

[2] Department of Human Pathology and Oncology, University of Florence, Florence, Italy.

[3] Gerhard-Domagk Institut fur Pathologie, Universitat von Münster, Münster, Germany.

[4] Department of Pathology, The Breast Unit, HELIOS Medical Center, Berlin, Germany.

[5] Laboratorio De Histopatologica, Centro Regional de Oncologia De Coimbra, Coimbra, Portugal.

[6] Center of Laboratory Medicine and Pathology, Department of Pathology, Rigshospitalet, Copenhagen University Hospital, Copenhagen, Denmark.

[7] Department of Histopathology, St. Bartholomew's Hospital Medical School, London, United Kingdom.

Supported by grant ETT 176/01 from the Hungarian Ministry of Health.

Gábor Cserni is supported by a János Bolyai Research Fellowship from the Hungarian Academy of Sciences.

This project would not have been possible without the organizers (Dr. Fritz Rank and Professors Simonetta Bianchi and Werner Boecker) or without the sponsors of the Copenhagen, Florence, and Münster meetings of the European Working Group for Breast Screening Pathology: Ethicon Endo-Surgery Breast Care (Italy), Leica Microsystems S.p.A. (Milan, Italy), and the German screening program.

The authors express special thanks to Béla Fekete (Kecskemét) and Eberhard Korsching (Münster) for arranging the web-based supplementary material.

BACKGROUND. The latest edition of the tumor-lymph node-metastasis (TNM) classification of malignant tumors distinguishes between isolated tumor cells (pN0) and micrometastases (pN1mi). The reproducibility of these categories has not been assessed previously.

METHODS. Digital images from 50 cases with low-volume lymph node involvement from axillary sentinel lymph nodes were circulated twice for evaluation (Evaluation Rounds 1 and 2) among the members of the European Working Group for Breast Screening Pathology, and the members were asked to categorize lesions as micrometastasis, isolated tumor cells, or something else and to classify each case into a pathologic lymph node (pN) category of the pathologic TNM system. Methods for improving the low reproducibility of the categorizations were discussed between the two evaluation rounds. κ Statistics were used for the assessment of interobserver variability.

The European Working Group for Breast Screening Pathology includes the following investigators: I. Amendoeira (Instituto de Patologia and Imunologia Molecular da Universidade do Porto e Serviço de Anatomia Patologica, Hospital S Joaõ, Porto, Portugal), N. Apostolikas (Department of Pathology, Saint Savas Hospital, Athens, Greece), J. P. Bellocq (Service d'Anatomie Pathologique, Hopital de Hautepierre, Strasbourg, France), S. Bianchi (Department of Human Pathology and Oncology, University of Florence, Florence, Italy), W. Boecker (Gerhard-Domagk Institut fur Pathologie, Universitat von Münster, Münster, Germany), B. Borisch (Department of Pathology Centre Médicale Universitaire, Genéve, Switzerland), C. E. Connolly (Department of Pathology, Clinical Sciences Institute, University College Hospital, Galway, Ireland), G. Cserni (Department of Pathology, Bács-Kiskun County Teaching Hospital, Kecskemét, Hungary), T. Decker (Department of Pathology, The Breast Unit, HELIOS Medical Center, Berlin, Germany), P. Dervan (Pathology Department, Mater Hospital, University College Dublin, Dublin, Ireland), M. Drijkoningen (Pathologische Ontleedkunde, University Hospital, Leuven, Belgium), I. O. Ellis (Department of Histopathology, City Hospital, Nottingham, United Kingdom), V. Eusebi (Sezione Anatomia Patologica M. Malpighi, Universita di Bologna, Ospedale Bellaria, Bologna, Italy), D. Faverly (CMP Laboratory, Bruxelles, Belgium), P. Heikkila (Haartman Institute, University of Helsinki, Helsinki, Finland), R. Holland (National Expert and Training Center for Breast Cancer Screening, Academisch Ziekenhuis Nijmegen, Nijmegen, The Netherlands), H. Kerner (Department of Pathology, Rambam Hospital, Haifa, Israel), J. Kulka (Second Department of Pathology, Semmelweis University, Budapest, Hungary), J. Jacquemier (Institut Paoli Calmettes, Marseille, France), M. Lacerda (Laboratorio De Histopatologica, Centro Regional De Oncologia De Coimbra, Coimbra, Portugal), J. Martinez-Penuela (Department of Pathology, Hospital de Navarra, Pamplona, Spain), C. De Miguel (Department of Pathology, Hospital Virgen del Camino, Pamplona, Spain), H. Nordgren (Patologilaboratoriet, Akademiska Sjukhuset, Uppsala, Sweden), J. L. Peterse (Department of Pathology, The Netherlands Cancer Institute, Amsterdam, The Netherlands), F. Rank (Center of Laboratory Medicine and Pathology, Department of Pathology, Rigshospitalet, Copenhagen University Hospital, Copenhagen, Denmark), P. Regitnig (Institute of Pathology, Medical University Graz, Graz, Austria), A. Reiner (Institute of Pathology, Donauspital, Wien, Austria), A. Sapino (Department of Biological Science and Human Oncology, University of Turin, Turin, Italy), B. Sigal-Zafrani (Institut Curie, Section Medicale et Hospitaliere, Paris, France), A. M. Tanous (Division d'Anatomie Pathologique, Laboratoire National de Santé, Luxembourg), S. Thorstenson (Department of Pathology and Cytology, Kalmar Hospital, Kalmar, Sweden), and C. A. Wells (Department of Histopathology, St. Bartholomew's Hospital Medical School, London, United Kingdom).

Address for reprints: Gábor Cserni, M.D., Ph.D., Bács-Kiskun County Teaching Hospital, Nyíri út 38., H-6000 Kecskemét, Hungary; Fax: (011) 36 76481219; E-mail: cserni@freemail.hu

Received July 7, 2004; revision received September 9, 2004; accepted September 21, 2004.

DOI 10.1002/cncr.20760
Published online 8 December 2004 in Wiley InterScience (www.interscience.wiley.com).

RESULTS. The κ value for the consistency of categorizing low-volume lymph node load into micrometastasis, isolated tumor cells, or neither of those changed from 0.39 to 0.49 between Evaluation Rounds 1 and 2, but it was slightly lower for the pN categories (0.35 and 0.44, respectively). Interpretation of the definitions of isolated tumor cells (especially with respect to their localization within the lymph node), lack of guidance on how to measure them if they were multiple, and lack of any definitions for multiple simultaneous foci of lymph node involvement were listed among the causes of discordant diagnoses.

CONCLUSIONS. The results of the current study indicated that the definitions available have minor contradictions and do not permit a reproducible distinction between micrometastases and isolated tumor cells. Refinement of these definitions, therefore, is required. One refinement that may improve reproducibility is suggested in this report. ***Cancer* 2005;103:358–67.**

KEYWORDS: immunohistochemistry, isolated tumor cells, kappa statistics, micrometastasis, sentinel lymph node, tumor-lymphnode-metastasis (TNM).

Lymph node status remains the most important single prognostic factor for patients with breast cancer, and it forms one of the main pillars of the tumor-lymph node-metastasis (TNM) classification system. Lymph node status traditionally has been determined by a single hematoxylin and eosin (H&E)-stained section from each lymph node removed by complete axillary dissection. Due to several factors, including breast screening programs, the prognostic profile for patients with breast cancer has improved, and the proportion of lymph node-positive tumors has decreased. Because patients with breast cancer derive no benefit from the removal of histologically negative lymph nodes, sentinel lymph node (SLN) biopsy, a minimally invasive staging procedure, is becoming a more common surgical method of axillary lymph node staging for patients with this disease. It not only allows SLN-negative patients to avoid lymph node clearance, but it also has been claimed that SLN biopsy improves staging accuracy by allowing a more cost-effective, concentrated, pathologic assessment of the lymph nodes.

The detection of occult metastases in lymph nodes was highlighted first over half a century ago,[1] but the clinical significance of these occult metastases, many of which belong to the category of micrometastasis or isolated tumor cells (ITCs), remains unsettled.[2] The widespread use of SLN biopsy has increased the number of patients with lymph node-positive breast cancer without any substantial change in the overall prognosis of the disease.[3] This stage migration and its potential misleading effects and artifacts in prognostication have highlighted a need to differentiate between metastases that may or may not be relevant. Micrometastases have been defined, according to the latest revision of the TNM classification system,[4] as metastases not $>$ than 2 mm, but $>$ 0.2 mm. A smaller volume lymph node load has been designated as ITCs, which are not $>$ 0.2 mm and are associated with qualitative features, such as the lack of metastatic activity or lack of extravascular/extrasinusoidal localization.[5] ITCs are not considered metastases and are regarded as pathologic lymph node negative (pN0), which is recorded as pN0(i+).[4–6]

The European Working Group for Breast Screening Pathology (EWGBSP) recently assessed the pathologic practice of reporting SLNs in Europe and found that the terms "micrometastasis" and "isolated tumor cells" were used quite heterogeneously.[7] EWGBSP members also felt that the definitions of these categories were not descriptive enough; therefore, a reproducibility study was initiated at our Copenhagen meeting in May, 2003. Here, we report the results of that study.

MATERIALS AND METHODS

Fifty low-volume lymph node metastases or ITCs (including 1 metastasis that measured $>$ 2 mm, 2 cases of capsular nevus, and 1 weakly cytokeratin-positive cell in the subcapsular sinus that was believed to represent a histiocyte) were selected from SLN biopsy material of the Bács-Kiskun County Hospital and the University of Coimbra. These materials were stained either with H&E or with immunohistochemistry (IHC) against cytokeratin. One-megapixel digital images were taken of each case with a conventional digital camera (Olympus Camedia 4040) attached to a conventional microscope. One to four images at different magnifications (generally medium-power and high-power views) of each lesion were captured. Images were saved as .jpg files on CD-ROMs and were sent to members of the EWGBSP. Each filename contained the case identifi-

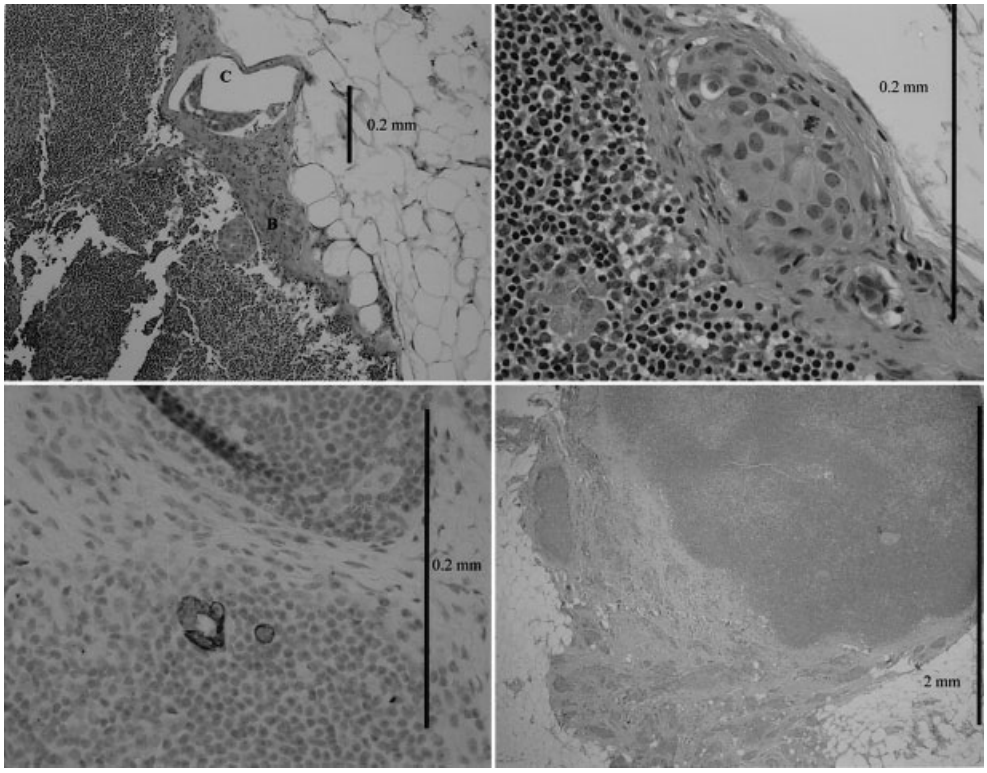

FIGURE 1. Illustrative cases from the study material. Images 2AM (top left), 5BH (top right), 19BH (bottom left), and 48AL (bottom right) from Cases 2, 5, 19, and 48, respectively (all sections were stained with hematoxylin and eosin [H&E] except 19BH, which was stained with immunohistochemistry against cytokeratin). For details of image labels, see Materials and Methods. Case 2, with its 2 foci, was upgraded from a majority diagnosis of isolated tumor cells (ITCs) (Evaluation Round 1) to a majority diagnosis of micrometastasis (Evaluation Round 2) because of the $>$ 0.2-mm size cluster in the capsular lymphatic vessel. Case 5 had a weak majority diagnosis of ITC, which was changed to equal ratings of ITC and micrometastasis; proliferation (a mitosis also is shown) and contact with lymphatic vessel walls would favor the diagnosis of micrometastasis, despite the $<$ 0.2-mm size. Case 19 had a unanimous categorization as ITC (Evaluation Round 1), and this was changed to a weak majority diagnosis of micrometastasis, because 52% of the responders felt that the single cell and the tubule were located within the parenchyma. The proportion of responders who considered this lesion as a single focus of lymph node involvement changed from 0.39 to 0.68. Case 48 represents the only metastasis that measured $>$ 2 mm that involved only the capsule and the extranodal fat. Although most responders felt that this was a (macro)metastasis, 5 responders did not classify this lesion as pN1, because they felt that the metastasis was not in the lymph node itself.

cation number (Cases 1–50), image identification (letters A–D, as appropriate, depending on the number of images on each case), and the magnification at which the image was taken (low power [L], $\times$ 40 magnification; medium-power [M], $\times$ 100 magnification; or high-power [H], $\times$ 400 magnification). All cases had a black scale bar representing either 0.2 mm or 2 mm (Fig. 1). The images used for the study can be downloaded from http://www.skyline-computer.hu/csernig.zip and are available on line at http://breast-pathology.uni-muenster.de/ (select "News from the EWGBSP" and supplementary material for this publication).

Members of the EWGBSP were asked to report each lesion independently by filling in a previously created Microsoft Excel spreadsheet. Members were asked to categorize each lesion (representing a row in the spreadsheet with identification numbers in column A) as micrometastasis (MIC), ITC, or something else (OTH) (column B). Whenever OTH was used as category, participants were asked to specify what they thought the lesion depicted was (column C); they also had to specify whether they thought the lesion was a single lesion or represented multiple foci of lymph node involvement (column D) and, where multiple, to state the number of foci identified (column E). Members were asked to assign a pN category according to the 6th edition of the *TNM Classification of Malignant Tumors*[4,8] (column F). Finally, the participants could give a yes or no answer on whether or not they were confident in their classification (column G). Comments were allowed in the last column (H).

After collecting the responses, κ statistics were calculated to evaluate the reproducibility of diagnosing and reporting micrometastases.[9] The advantage of these statistics is that values of κ are independent of

the real category: They simply reflect the consistency of ratings by the observers. The following arbitrary limits and labels were used to interpret agreement on the basis of the κ values: $<$ 0.00, poor; 0.00–0.20, slight; 0.21–0.40, fair; 0.41–0.60, moderate; 0.61–0.80, substantial; and 0.81–1.00, very good.[10] Only the κ values were shared with members of the EWGBSP without analyzing the distribution of majority diagnoses or categorizations and without discussing any of the cases. The possible causes of the low reproducibility were discussed at a meeting of the EWGBSP in Florence, Italy (January, 2004), and a document was drafted to improve the reproducibility of differentiating between micrometastases and ITCs. This document was discussed further and was finalized after consultation with experts of the International Committee Against Cancer (UICC) TNM committee (Drs. Christian Wittekind and Leslie Sobin), and members of the EWGBSP were asked to go through all of the cases again and use the document to categorize the lesions. κ Scores were calculated again, and the results were discussed at the next meeting of the EWGBSP in Münster, Germany (April, 2004). At that meeting, the majority diagnoses were also shared, and a case-by-case discussion was held.

RESULTS

Evaluation Round 1

Evaluation Round 1 was completed by 24 members of the EWGBSP, but only 23 members provided a pN classification. There were only two cases that were classified unanimously as MIC, ITC, or OTH, and only one case received the same pN category. The majority classifications of interest of this evaluation round, as well as the confidence ratings, are shown in columns B, D, and F of Table 1, whereas the κ scores for the given categories of interest are presented in Table 2. Overall, the reproducibility of the categorization into groups of MIC, ITC, or OTH as well as into the pN categories was fair, and this was considered suboptimal by the EWGBSP. Reproducibility of diagnosing micrometastases was somewhat better and fell in the moderate range.

Refinements of the Descriptions of ITC and Micrometastases

On the basis of recent publications,[4–6] considerations listed below (see Discussion), and consultation with experts from the TNM project, we decided to adopt the definition of ITCs from the 6th edition of the *TNM Classification of Malignant Tumors*.[4] Suggestions were made on how to measure the size of the ITC or metastasis in a lymph node and on the determination of the number of tumor foci: 1) If there are multiple foci, then only the largest should be considered. (The rationale for this stems from General Staging Rules No. 4 and 5 of the TNM classification system, i.e., if there is doubt, then the lower category should be chosen, and the same approach is used for multiple primary tumors, respectively.) For what should be considered a single focus and what should be considered multiple foci, see below. 2) If single tumor cells, clusters, or nests are continuous or are separated by a distance of a few cells (e.g., two to five cells), then consider them as one focus and measure the largest size. 3) If the cells, clusters, or nests are discontinuous and are dispersed homogeneously (evenly; e.g., multiple tubules or lobular carcinoma cells separated by lymphoid tissue but occupying a definable part of the lymph node, even if this is the whole lymph node), then measure them as one. (The rationale for this is that these patterns represent a diffuse, noncohesive involvement of the lymph node, and it is more logical to suggest that all the volume affected is involved rather than trying to reduce the extent of lymph node involvement.) 4) If the cells, clusters, or nests are discontinuous and are dispersed unevenly (e.g., two or a few cells or clusters at some distance from each other), then consider it as one if the distance between foci is smaller than the smaller cluster. Otherwise, if the distance between unevenly distributed tumor cells, clusters, or nests is greater, then consider it as two or more foci and measure the largest cluster. (The rationale for this is that two or more foci, when distant enough, are seen as two or more; however, when they are close to one another, they seem to merge, and the chances of missing an unsampled junction between the two or more foci increase.)

Suggestions also were made to include the location of the tumor cells in the definition. It was stressed that the capsule is part of the lymph node; therefore, tumor cells that are in the capsule (including lymphatic vessels) are in the lymph node. Consequently, if something grows in the capsule and outside of it, then this should be considered as lymph node involvement. An analogy supporting this decision was found in the recent TNM supplement,[11] which stated that, in defining the N classification, the perinodal component should be included in the size of the isolated lymph node metastasis.

It also was agreed that tumor cells (single or in clusters) clearly in the parenchyma (and not in the sinuses or vascular spaces) should be categorized as micrometastasis (pN1mi), even if they measured $<$ 0.2 mm and had no associated proliferation or stromal reaction. (Signs of proliferation or stromal reaction rarely are associated with low-volume metastases.) If tumor cells (single or in clusters) are localized in ves-

TABLE 1
Categorization of the Cases in the Two Evaluation Rounds

	Majority category (proportion of reports)		Majority pN classification (proportion of reports)		Proportion of EWGBSP members who were confident about their classification	
A. Case no. (no. of images)	B. Round 1 (n = 24 reports)	C. Round 2 (n = 25 reports)	D. Round 1 (n = 23 reports)	E. Round 2 (n = 25 reports)	F. Round 1	G. Round 2
C1 (2)	MIC (0.88)	MIC (1.00)	pN1mi (0.87)	pN1mi (1.00)	1.00	1.00
C2 (3)	ITC (0.71)	MIC (0.72)	pN0 (0.39)	pN1mi (0.72)	0.63	0.71
C3 (2)	MIC (0.92)	MIC (1.00)	pN1mi (0.91)	pN1mi (1.00)	0.92	1.00
C4 (2)	ITC (0.71)	ITC (0.96)	pN0 (0.65)	pN0(i+) (0.84)	0.83	0.92
C5 (2)	ITC (0.50)	MIC/ITC (0.48)	pN0 (0.43)	pN1mi (0.48)	0.67	0.58
C6 (2)	MIC (0.79)	MIC (0.96)	pN1mi (0.83)	pN1mi (0.94)	0.83	0.83
C7 (2)	ITC (0.92)	ITC (0.64)	pN0(i+) (0.83)	pN0(i+) (0.64)	0.79	0.63
C8 (2)	ITC (0.92)	ITC (0.64)	pN0(i+) (0.78)	pN0(i+) (0.64)	0.83	0.71
C9 (2)	ITC (0.63)	MIC (0.68)	pN0 (0.48)	pN1mi (0.68)	0.67	0.63
C10 (2)	ITC (0.88)	ITC (0.96)	pN0 (0.48)	pN0(i+) (0.96)	0.88	0.96
C11 (2)	OTH (0.50)	MIC (0.76)	pN0 (0.48)	pN1mi (0.80)	0.63	0.79
C12 (2)	MIC (0.54)	MIC (0.56)	pN1mi (0.52)	pN1mi (0.52)	0.58	0.79
C13 (3)	MIC (0.50)	MIC (0.84)	pN1mi (0.52)	pN1mi (0.84)	0.63	0.58
C14 (2)	MIC (0.71)	MIC (0.84)	pN1mi (0.70)	pN1mi (0.84)	0.67	0.83
C15 (2)	MIC (0.50)	MIC (0.80)	pN1mi (0.48)	pN1mi (0.80)	0.46	0.71
C16 (2)	MIC (0.96)	MIC (1.00)	pN1mi (0.96)	pN1mi (1.00)	0.92	1.00
C17 (2)	MIC/OTH (0.46)	MIC (0.84)	pN0 (0.48)	pN1mi (0.84)	0.38	0.63
C18 (1)	MIC (0.88)	MIC (1.00)	pN1mi (0.87)	pN1mi (0.96)	0.88	1.00
C19 (2)	ITC (1.00)	MIC (0.52)	pN0(i+) (0.91)	pN1mi (0.52)	0.92	0.71
C20 (2)	ITC (0.75)	MIC (0.68)	pN0(i+) (0.70)	pN1mi (0.72)	0.75	0.63
C21 (2)	ITC (0.83)	MIC (0.64)	pN0(i+) (0.83)	pN1mi (0.64)	0.92	0.75
C22 (2)	OTH (0.58)	OTH (0.56)	pN0 (0.61)	pN0 (0.56)	0.63	0.42
C23 (3)	ITC (0.92)	MIC (0.60)	pN0(i+) (0.87)	pN1mi (0.60)	0.88	0.54
C24 (3)	ITC (0.71)	ITC (0.56)	pN0(i+) (0.70)	pN0(i+) (0.56)	0.63	0.50
C25 (1)	ITC (0.75)	ITC (0.88)	pN0(i+) (0.78)	pN0(i+) (0.88)	0.54	0.54
C26 (1)	ITC (0.58)	ITC (0.68)	pN0(i+) (0.65)	pN0(i+) (0.68)	0.67	0.58
C27 (2)	ITC (0.71)	ITC (0.88)	pN0(i+) (0.70)	pN0(i+) (0.88)	0.50	0.50
C28 (4)	ITC (0.54)	MIC (0.80)	pN1mi (0.39)	pN1mi (0.80)	0.50	0.75
C29 (2)	ITC (0.67)	ITC (0.80)	pN0 (0.48)	pN0(i+) (0.76)	0.71	0.46
C30 (3)	MIC (0.75)	MIC (0.96)	pN1mi (0.74)	pN1mi (0.96)	0.58	0.67
C31 (2)	ITC (0.63)	MIC (0.68)	pN1mi (0.39)	pN1mi (0.68)	0.54	0.71
C32 (2)	ITC (0.96)	ITC (0.92)	pN0 (0.52)	pN0(i+) (0.84)	0.88	0.92
C33 (1)	ITC (0.58)	ITC (0.96)	pN0 (0.70)	pN0(i+) (0.88)	0.71	0.92
C34 (2)	ITC (0.63)	ITC (0.84)	pN0 (0.74)	pN0(i+) (0.72)	0.71	0.88
C35 (2)	MIC (0.92)	MIC (1.00)	pN1mi (0.91)	pN1mi (1.00)	0.96	0.96
C36 (1)	MIC (0.83)	MIC (1.00)	pN1mi (0.83)	pN1mi (1.00)	0.83	1.00
C37 (2)	ITC (0.54)	ITC (0.88)	pN0 (0.70)	pN0(i+) (0.76)	0.71	0.58
C38 (2)	ITC (0.63)	ITC (0.96)	pN0 (0.70)	pN0(i+) (0.88)	0.88	0.92
C39 (1)	ITC (1.00)	ITC (0.60)	pN0(i+) (0.87)	pN0(i+) (0.60)	0.92	0.71
C40 (1)	ITC (0.83)	ITC (0.56)	pN0(i+) (0.65)	pN0(i+) (0.56)	0.63	0.54
C41 (2)	OTH (0.63)	MIC (0.64)	pN0 (0.57)	pN1mi (0.64)	0.54	0.58
C42 (3)	MIC (0.63)	MIC (0.84)	pN1mi (0.48)	pN1mi (0.84)	0.50	0.79
C43 (4)	MIC (0.67)	MIC (0.96)	pN1mi (0.65)	pN1mi (0.96)	0.58	0.79
C44 (2)	MIC (0.92)	MIC (1.00)	pN1mi (0.96)	pN1mi (1.00)	0.88	1.00
C45 (2)	MIC (0.88)	MIC (1.00)	pN1mi (0.91)	pN1mi (1.00)	0.92	1.00
C46 (2)	ITC (0.92)	ITC (0.52)	pN0(i+)/pN0 (0.43)	pN1mi (0.48)	0.71	0.54
C47 (3)	MIC (0.83)	MIC (1.00)	pN1mi (0.78)	pN1mi (1.00)	0.67	0.83
C48 (2)	OTH (0.83)	OTH (0.92)	pN1a (0.78)	pN1a (0.92)	0.79	0.88
C49 (1)	ITC (0.50)	ITC (0.68)	pN0(i+)/pN0 (0.48)	pN0(i+) (0.68)	0.54	0.58
C50 (4)	MIC (0.67)	MIC (0.84)	pN1mi (0.52)	pN1mi (0.76)	0.79	0.71

pN: pathologic lymph node status; EWGBSP: European Working Group for Breast Screening Pathology; MIC, mi: micrometastases; ITC, i+: isolated tumor cells; OTH: other.

TABLE 2
κ Scores for Diagnostic and Staging Categories in the Two Evaluation Rounds

	Round 1		Round 2	
κ Score	A. All cases (n = 50)	B. MIC and ITC only (n = 46)	C. All cases (n = 50)	D. MIC and ITC only (n = 46)
MIC	0.45	0.47	0.53	0.53
ITC	0.42	0.41	0.48	0.48
OTH	0.23	0.14	0.35	0.08
SE of κ score (MIC, ITC, OTH)	±0.009	±0.009	±0.008	±0.009
Overall κ score	0.39	0.38	0.49	0.48
SE of overall κ score	±0.007	±0.007	±0.007	±0.008
pN0 κ score	0.22	0.23	0.15	0.15
pN0(i+) κ score	0.35	0.34	0.42	0.46
pN1mi κ score	0.46	0.47	0.53	0.57
pN1a κ score	0.40	0.19	0.74	0.17
SE of κ score (pN categories)	±0.009	±0.009	±0.008	±0.008
Overall κ score	0.35	0.35	0.45	0.32
SE of overall κ score	±0.006	±0.006	±0.006	±0.007

MIC, mi: micrometastases; ITC, i+: isolated tumor cells; OTH: other; SE: standard error; pN: pathologic lymph node status.

sels or sinuses, then it was agreed that the size was to be used as a defining measure, with not $>$ 0.2 mm considered ITC (pN0[i+]) and with $>$ 0.2 mm but not $>$ 2 mm considered micrometastasis (pN1mi).

Evaluation Round 2 was carried out using these guidelines. The full document used for the second evaluation round can be downloaded from http://www.skyline-computer.hu/csernig.zip (filename: newreadme.rtf) and is available on line at http://breast-pathology.uni-muenster.de/ (select "News from the EWGBSP" and supplementary material for this publication; this text is in "Definitions, procedures and categories (pdf)."

Evaluation Round 2

Evaluation Round 2 was completed by 25 members of the EWGBSP. There were 9 cases that were classified unanimously as MIC, ITC, or other (OTH) and 8 cases that received the same pN category, which was a considerable improvement over Evaluation Round 1. The majority classifications of interest of this round, as well as the confidence ratings are shown in columns C, E, and G of Table 1; whereas the κ scores for the given categories of interest are presented in Table 2. Overall, the reproducibility of the categorization into groups of MIC, ITC, or OTH as well as into the pN categories improved and fell into the moderate range. The reproducibility of diagnosing both micrometastases and ITCs improved in Evaluation Round 2 but still fell into the moderate range.

DISCUSSION

There are many aspects of SLN biopsy that are different from institution to institution, including the quality and quantity of the tracers, the method and timing of their administration, the definition of a radioactive lymph node, the handling, and the histologic investigation of the SLNs.[2] Interpretation of the pathologic findings is probably the most ignored aspect of the differences between laboratories. A recent publication pointed to the possible interobserver differences in distinguishing positive and negative SLNs by circulating a single set of 25 H&E-stained and IHC-stained SLNs between 10 pathologists.[12] Only three of those cases were reported unanimously, and low-volume metastases, especially those that were demonstrated by cytokeratin IHC, were called negative more often. However, it is not known how much these results were dependent either on the detection of lymph node involvement or its lack; or on the interpretation of ITCs composed of one or few cells as negative or positive findings; or, finally, on the distinction between ITCs and micrometastases. The cited article[12] did not distinguish between ITCs and micrometastasis in its terminology and reflected the earlier edition of the TNM classification (with micrometastasis defined as anything not $>$ 2 mm),[13,14] but some participants may have been using the newer edition[4,6,8] for the pN0–pN1 categories. It is obvious from two studies that used automated technologies that the detection of low-volume lymph node involvement by conventional microscopy is not perfect and is compromised by factors such as human fatigue or expertise.[15,16] Observers in our study, all with special expertise in breast pathology, did not have to search for the foci of lymph node involvement but only had to interpret the depicted lesion; therefore the current study concentrated on interpretation issues together with the reproducibility of categorizing these lesions according to the pN categories from the latest version of the TNM classification system.

The TNM system is an accepted standard for describing the anatomic extent of disease in cancer. It helps in planning treatments, in prognostication, in evaluating the results of different treatment options, and in communicating between different treatment centers. It also contributes to the continuing investigation of human malignancies.[17,18] It has been accepted as the recommended staging system for breast cancer in the World Health Organization classification of breast tumors[19] and has been incorporated in the 4th edition of the European Guidelines, which is in preparation.[20] Because of these considerations, we be-

lieved it was important to check the reproducibility of reporting low-volume lymph node involvement.

The poor reproducibility of the classification into MIC, ITC, or OTH and the different pN categories was discussed by the members of the EWGBSP after Evaluation Round 1, and the following factors were identified as possible contributors: 1) The definitions of ITCs (this newly introduced category of "nonmetastatic" lymph node involvement) are somewhat heterogeneous and contradictory. The first definition of this category,[5] which was cited in further definitions,[4,8] includes a table suggesting that ITCs have no contact with vessel or lymph sinus wall, are not located outside of the vessel or lymph sinus wall, and have no extravascular or extrasinusoidal stromal reaction or proliferation. For example, this table would enable tumor cell clusters measuring < 0.2 mm to be labeled as micrometastasis if they showed any of the features ITCs do not show (e.g., extrasinusoidal [parenchymal] localization). This also is suggested by the wording of the definition for ITCs as single tumor cells or small clusters of cells not > 0.2 mm in greatest dimension that usually are detected by immunohistochemistry or molecular methods but that may be verified with H&E staining. ITCs *typically* do not show evidence of metastatic activity (e.g., proliferation, stromal reaction) or penetration of vascular or lymphatic sinus walls.[4] However, the definition of micrometastasis (> 0.2 mm, but none > 2 mm in greatest dimension[4]) somewhat contradicts this by stating the minimum size a micrometastasis should reach to be labeled as such. ITCs identified by morphologic methods (i.e., microscopy) should be labeled pN0(i+), according to their definition as no regional lymph node metastasis histologically and positive morphologic findings for ITC,[4] with the first part referring to the fact that ITCs are not considered metastases the second part referring to their detection by microscopy.

The definitions given in the American Joint Committee on Cancer (AJCC) *Cancer Staging Handbook*[8] show minor differences with the definitions cited above: ITCs are defined as single tumor cells or small clusters of cells not > 0.2 mm in greatest dimension, usually with no malignant activity (such as proliferation or stromal reaction). If an additional immunohistochemical examination is made for ITCs in a patient with histologically negative lymph nodes, then the regional lymph nodes should be designated as either pN0(i−) or pN0(i+), as appropriate. According to this latter definition, which lacks localization as a defining criterion and uses the word *usually* instead of *typically* (typical defines the type, whereas usual is common but not specific), pN0(i+) was defined as no regional lymph node metastasis histologically, positive IHC, and no IHC cluster measuring > 0.2 mm; and the i+ refers to IHC (a method of detection) and not ITCs (the entity detected). This difference obviously was responsible in part for the lower κ scores for the pN categories. The AJCC definition was modified between the two evaluation rounds,[6] and now a uniform use and interpretation for the (i+) symbol is ITC. This change also was considered in the second round of evaluation.

Another TNM-related publication also may have been a source of differing interpretation: In that report, micrometastases are distinguished from ITCs on the basis of size, and they are more likely to show histologic evidence of microscopic malignant activity, although this is not an absolute requirement. A critical element of this definition is that the distinction between micrometastases and ITCs is made based on size alone.[21] Although we agree that this probably would allow the most reproducible categorization of micrometastases and ITCs, current definitions in use[4,8] do not fully support this approach.

Another source of potential confusion for general pathologists may be the fact that ITC, as defined above, means more than the meaning of its three words *isolated tumor cells* and also includes small clusters of cohesive, hence *nonisolated*, tumor cells. This is why some researchers have preferred the term *submicrometastasis* to ITCs to describe the same entity.[22,23]

2) In considering the size of the lesion, it is critical to decide whether to regard the lesion as a single larger focus or as several smaller foci; i.e., a micrometastasis versus several ITCs or a metastasis versus multiple micrometastases. No guidance for this decision was available to us.

3) The localization of the lesions also was interpreted differently and resulted in confusing classifications from several aspects. Interpretations using the UICC publications[4,5] considered some of the intraparenchymal lesions measuring < 0.2 mm as micrometastasis because of their extrasinusoidal manifestation, whereas others used size alone as a defining criterion.

Some lesions represented tumor cells or small clusters in the capsular lymphatic channels, and some observers interpreted these not as lymph node involvement but as lymphatic invasion and, consequently, ranked the lesions as OTH and categorized them as pN0. Two cases (1 measuring > 2 mm) were from capsular and extracapsular involvement without spread to the subcapsular area or parenchyma of the SLN. Some interpreters did not classify these lesions as lymph node involvement but commented on extranodal axillary fat involvement.

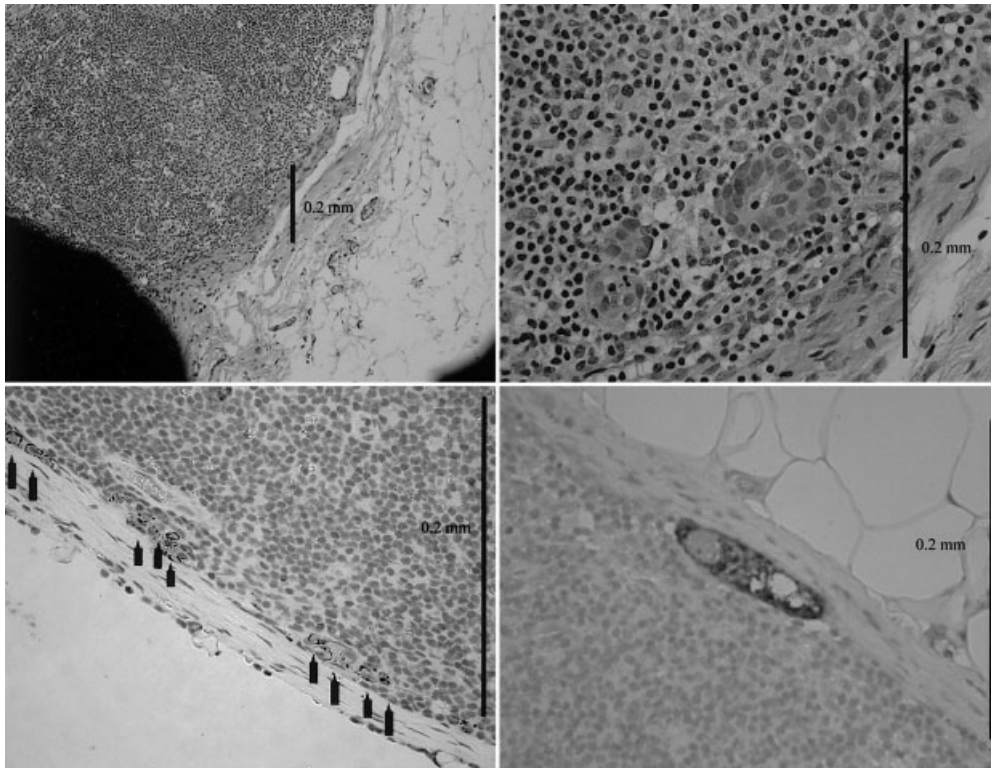

FIGURE 2. Cases with a significant change in interpretation. Images 46AM (top left), 46BH (top right), 40AH (bottom left), and 49AH (bottom right) from Cases 46, 40, and 49, respectively (Case 46 was stained with hematoxylin and eosin; Cases 40 and 49 were stained with immunochemistry against cytokeratin). For details of image labels, see Materials and Methods. For details about the changes in interpretation, see Discussion. Note that the markers on the suboptimally stained, cytokeratin-positive cells in the bottom left image were missing from the study material.

4) The test was performed on digital images, resulting in a lack of comparison with cells of the primary tumor and differences between the images and the real microscopic views. It is probable that this fact also may have reduced interobserver agreement.

5) Because the objective of the test was to assess the reproducibility of categorizing lymph node involvement, the inclusion of the OTH category lesions (2 capsular nevi, 1 macrophage, and a metastasis measuring > 2 mm) also was disturbing. This was seen especially in Evaluation Round 2. Because it was felt that these cases really were not problematic for recognition on real slides and were not related fully to the test, κ statistics also were calculated after removal of these four cases, although their elimination resulted in a negligible increase in the κ values for micrometastases and ITCs. In parallel, the κ values for OTH (pN0 and pN1a categories) decreased substantially, because there were no cases left with these as the original diagnosis, and there remained only a few readings in these categories (Table 2, compare columns B and D with columns A and C, respectively). These factors all were discussed and were included in the consensus document that was used to improve the diagnostic accuracy of ITCs and micrometastases in Evaluation Round 2.

It must be mentioned that a few cases (e.g., Cases 4, 29, and 46) were labeled ITC and were categorized as pN0(i+) in a somewhat different proportion during Evaluation Round 2 (Table 1). The same phenomenon can be seen with MIC and pN1mi (e.g., Cases 6, 11, and 50). This is simply the result of some observers inconsistently neglecting to add the qualifiers "i+" or "mi" to the main pN category. Although it is clear that the identification of ITC and MIC should lead to categorization as pN0(i+) and pN1mi, respectively, the human phenomenon of not fully following written guidelines should not be forgotten whenever scoring subjective human performances, such as ratings into different categories.

A few cases required special attention because of considerable differences in categorization between Evaluation Rounds 1 and 2 (Table 1). The single most important contributing factor responsible for these differences was the localization of tumor cells or clusters measuring < 0.2 mm within the parenchyma. This resulted in the upstaging of Cases 2 (Fig. 1), 9, 19 (Fig. 1), 20, 21, 23, 28, and 31 from ITC to MIC. This same factor accounted for the smaller number of ratings into ITC in Case 46, in which some observers interpreted that tumor cell clusters measuring < 0.2 mm were in the parenchyma, but this is arguable; therefore, ITC seems a good option both from the point of view of a probable sinusoidal localization and a degree of uncertainty (General Staging Rule No. 4) (Fig. 2).

The number of ratings into ITC in Case 40 decreased because of another factor, namely, the interpretation of multiple single cells and small clusters as a single micrometastasis instead of multiple foci of ITC (Fig. 2); whereas, in Case 49, the number of ratings into ITC increased, because the single cytokeratin-positive cell was interpreted by more observers as a tumor cell (Fig. 2).

By providing more precision to the definitions of ITC and micrometastases and solving some of the conflicting aspects of these definitions, the EWGBSP was able to improve the consistency of the diagnoses of these two categories. This improvement, however, resulted in a higher rate of micrometastasis diagnosed in cases that were diagnosed confidently as ITC in the first round. The ITC category is defined arbitrarily, and its upper size limit also is not evidence-based but can be estimated easily as a proportion of the high-power field area diameter in most microscopes. Because the ITC category was introduced to overcome the detailed SLN analysis-related stage migration artifact with all its potential harm in treatment decisions and cancer endpoint statistics,[3] it is not known whether the improved consistency achieved by our group of pathologists specialized in breast diseases is not acting against this objective. Strong evidence for the prognostic significance of micrometastases in breast cancer is lacking,[2] and there is even less evidence for any prognostic role for ITC, which clearly justifies the position of ITC in the pN0 class. The distinction between pN0 and pN1 often is considered critical for treatment decisions and also is critical from a medicolegal viewpoint relating to giving or not giving a specific treatment; therefore it is believed that the distinction between ITC/pN0(i+) and micrometastasis/pN1mi is important for staging purposes.

It should be noted that there always will be some cases that are difficult to classify and that pathologists always will have some doubts about their classifications, as shown in columns F and G on Table 1. There, General Staging Rule No. 4 of the TNM classification system should be the main help, and choosing the lower category is the best option.

In the current study, we showed that the definitions available (including minor contradictions in them) do not permit a reproducible distinction between the two entities of micrometastasis and ITC, and refinements of the definitions are required. After consultation with TNM experts, we offer one improvement in the definitions that we have shown to improve reproducibility. We admit that these refinements also are arbitrary; however, without reproducible diagnoses, there always will be a large bias in studies that try to assess the prognostic role of ITC or micrometastases.

REFERENCES

1. Pickren JW. Significance of occult metastases. A study of breast cancer. *Cancer.* 1961;14:1266–1271.
2. Cserni G, Amendoeira I, Apostolikas N, et al. Pathological work-up of sentinel lymph nodes in breast cancer. Review of current data to be considered for the formulation of guidelines. *Eur J Cancer.* 2003;39:1654–1667.
3. Weaver DL. Sentinel lymph nodes and breast carcinoma: which micrometastases are clinically significant? *Am J Surg Pathol.* 2003;27:842–845.
4. Sobin LH, Wittekind C, editors. UICC TNM classification of malignant tumours, 6th ed. New York: John Wiley & Sons, Inc., 2002.
5. Hermanek P, Hutter RVP, Sobin LH, Wittekind C. Classification of isolated tumor cells and micrometastasis. *Cancer.* 1999;86:2668–2673.
6. Singletary SE, Greene FL, Sobin LH. Classification of isolated tumor cells: clarification of the 6th edition of the American Joint Committee on Cancer staging manual. *Cancer.* 2003; 98:2740–2741.
7. Cserni G, Amendoeira I, Apostolikas N, et al. Discrepancies in current practice of pathological evaluation of sentinel lymph nodes in breast cancer. Results of a questionnaire-based survey by the European Working Group for Breast Screening Pathology. *J Clin Pathol.* 2004;57:695–701.
8. Greene FL, Page DL, Fleming ID, et al., editors. AJCC cancer staging handbook—TNM classification of malignant tumors, 6th ed. New York: Springer Verlag, 2002.
9. Fleiss JL. The measurement of interrater agreement. In: Fleiss JL, editor. Statistical methods for rates and proportions, 2nd ed. New York: John Wiley & Sons, Inc., 1981:212–236.
10. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics.* 1977;33:159–174.
11. Wittekind C, Greene FL, Henson DE, Hutter RVP, Sobin LH, editors. TNM supplement: a commentary on uniform use, 3rd ed. New York: John Wiley & Sons, Inc., 2003.
12. Roberts CA, Beitsch PD, Litz CE, et al. Interpretive disparity among pathologists in breast sentinel lymph node evaluation. *Am J Surg.* 2003;186:324–329.
13. Sobin LH, Wittekind C, editors. UICC TNM classification of malignant tumours, 5th ed. New York: John Wiley & Sons, Inc., 1997.
14. Fleming ID, Cooper JS, Henson DE, et al., editors. AJCC cancer staging manual, 5th ed. Philadelphia: Lippincott-Raven Publishers, 1997.
15. Weaver DL, Krag DN, Manna EA, Ashikaga T, Harlow SP, Bauer KD. Comparison of pathologist-detected and automated computer-assisted image analysis detected sentinel lymph node micrometastases in breast cancer. *Mod Pathol.* 2003;16:1159–1163.
16. Mesker W, Torrenga H, Sloos W, et al. Supervised automated microscopy increases sensitivity and efficiency of detection of sentinel node micrometastases in breast cancer patients. *J Clin Pathol.* 2004;57:960–964.
17. Gospodarowicz MK, Miller D, Groome PA, Greene FL, Logan PA, Sobin LH. The process for continuous improvement of the TNM classification. *Cancer.* 2004;100:1–5.
18. Greene FL, Sobin LH. The TNM system: our language for cancer care. *J Surg Oncol.* 2002;80:119–120.

19. Tavassoli FA, Devilee P, editors. World Health Organization Classification of tumours. Tumours of the breast and female genital organs. Lyon: IACR Press, 2003.
20. Perry N, Broeders M, de Wolf C, et al., editors. European guidelines for quality assurance in mammography screening, 3rd ed. Luxemburg: European Communities, 2001.
21. Singletary SE, Allred C, Ashley P, et al. Staging system for breast cancer: revisions for the 6th edition of the AJCC Cancer Staging Manual. *Surg Clin North Am.* 2003;83:803–819.
22. Schwartz GF, Giuliano AE, Veronesi U, Consensus Conference Committee. Proceedings of the consensus conference on the role of sentinel lymph node biopsy in carcinoma of the breast, April 19–22, 2001, Philadelphia, Pennsylvania. *Cancer.* 2002;94:2542–2551.
23. Schwartz GF, Giuliano AE, Veronesi U, Consensus Conference Committee. Proceedings of the consensus conference on the role of sentinel lymph node biopsy in carcinoma of the breast, April 19–22, 2001, Philadelphia, Pennsylvania. *Hum Pathol.* 2002;33:579–589.